



# Approche probabiliste des liens entre distances et maillages, application à l'exploitation d'enquêtes origine destination

Olivier Bonin, Mohamed Hamza Lemssouguer

## ► To cite this version:

Olivier Bonin, Mohamed Hamza Lemssouguer. Approche probabiliste des liens entre distances et maillages, application à l'exploitation d'enquêtes origine destination. Journées de méthodologie statistique, Jan 2012, France. pp.1. halshs-00737380

**HAL Id: halshs-00737380**

**<https://shs.hal.science/halshs-00737380>**

Submitted on 9 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **APPROCHE PROBABILISTE DES LIENS ENTRE DISTANCES ET MAILLAGES, APPLICATION A L'EXPLOITATION D'ENQUÊTES ORIGINE DESTINATION**

*Olivier BONIN(\*), Mohamed Hamza LEMSSOUGUER(\*\*)*

*(\*) UPE – IFSTTAR – LVMT*

*(\*\*) Ecole des Ponts Paristech*

## **Introduction**

Les enquêtes dites origine-destination (OD), et plus généralement les enquêtes ménages déplacements (EMD), décrivent des déplacements en s'appuyant sur un zonage qui est soit en lien avec un découpage administratif, soit constitué d'un carroyage (cas de l'EGT – enquête globale de transports, EMD de la région Ile-de-France). On connaît pour chaque déplacement sa zone de départ et sa zone d'arrivée. Cependant les enquêtes OD, ainsi que les EMD, ne recueillent pas de données chiffrées sur les longueurs des déplacements. Il est alors nécessaire d'estimer les distances parcourues lors des déplacements à partir des données des zones de départ, des zones d'arrivée, et éventuellement des temps de parcours.

Pour un déplacement, ce problème se décompose en trois points : détermination du lieu de départ à l'intérieur de la zone de départ, détermination du lieu d'arrivée à l'intérieur de la zone d'arrivée, et détermination de l'itinéraire emprunté et ainsi de la distance parcourue. Ce dernier point fait l'objet d'une note méthodologique du CERTU [CERTU], qui propose diverses formules pour estimer les distances réellement parcourues à partir des distances à vol d'oiseau entre centroïdes de zones. Cependant, le problème de la détermination du point de départ et du point d'arrivée du déplacement ne peut être abordé que de façon probabiliste, en tenant compte par exemple des distributions de population dans les zones de départ et d'arrivée. Nous évaluons dans cette communication l'impact de cette incertitude sur les distances estimées.

Nous commençons par la résolution de deux problèmes posés par Ch. Terrier [Terrier 09]. Ce sont des généralisations du problème de l'aiguille de Buffon, que nous étendons également au cas

des déplacements en ligne brisée par un calcul d'espérance faisant appel à un procédé de pseudo division euclidienne. Cette approche correspond au problème des enquêtes origine destination s'appuyant sur un carroyage régulier.

Nous abordons ensuite le problème des maillages irréguliers de la manière suivante : soit un déplacement effectué entre un point  $A$  et un point  $B$ , enregistré dans une enquête sous la forme d'un déplacement entre une zone origine  $o$  de centroïde  $C_o$  et une zone destination  $d$  de centroïde  $C_d$ . Nous relierons statistiquement la distance entre  $A$  et  $B$  à la distance entre  $C_o$  et  $C_d$ , étant données des hypothèses de distribution statistique de  $A$  autour de  $C_o$  et de  $B$  autour de  $C_d$ . Nous donnons une expression des lois de distance dans le cas général, et proposons des applications dans des cas particuliers (distribution Gaussienne autour des points de départ et d'arrivée, distribution uniforme dans les zones) qui permettent de mettre en évidence l'impact de la taille des zones de départ et d'arrivée, ainsi que de la concentration des origines et des destinations autour des centroïdes de zone.

## 1 Position du problème

Les déplacements renseignés dans les enquêtes origine-destination ne recueillant pas de données chiffrées sur les distances effectivement parcourues lors des déplacements : on ne dispose que de la zone de départ et de la zone d'arrivée.

Ces zones peuvent être de taille variable selon les enquêtes. Le recensement de la population fournit des données de ce type pour les déplacements domicile-travail à un niveau communal. Le maillage est donc administratif, avec des mailles taille très variable. Les EMD utilisent des maillages plus fins ; à titre d'exemple, l'EMD de Lyon comporte environ 700 zones. En Ile-de-France, l'EGT utilise un carroyage en guise maillage, avec des carreaux de 300 au plus fin.

Imputer une distance entre deux carreaux, ou entre deux zones, nécessite dans un premier temps de déterminer un point de départ et un point d'arrivée, qui sont soit des centroïdes des mailles, soit des points ayant un sens particulier (par exemple, les chefs-lieux de commune si le maillage est communal), puis de calculer la distance sur le réseau pour affiner la distance à vol d'oiseau. Le guide du CERTU [CERTU] a calibré les formules suivantes. En appelant  $D$  la vraie distance parcourue et  $D_{vo}$  la distance à vol d'oiseau, les formules calées sont pour des trajets en voiture et deux-roues motorisés :

- si  $D_{vo} \leq 1\text{km}$  alors  $D = (2,2 - 0,72D_{vo})D_{vo}$
- si  $D_{vo} \geq 1\text{km}$  alors  $D = 1,4D_{vo}$

pour les trajets en transport collectifs :

- $D = 1,5D_{vo}$

et pour les trajets en vélo :

$$- D = 1,35D_{vo}$$

On constate d'après ces formules que le facteur de correction est de l'ordre de 1,4, ce qui n'est pas étonnant puisque c'est le rapport entre la distance du trajet deux points d'un carré séparés par une diagonale et la distance du trajet pour rejoindre ces deux points en empruntant deux côtés (Figure 1), soit le rapport de la distance de Manhattan et de la distance euclidienne. La distance de Manhattan correspond intuitivement à des parcours de réseaux routiers en zone urbaine dense.

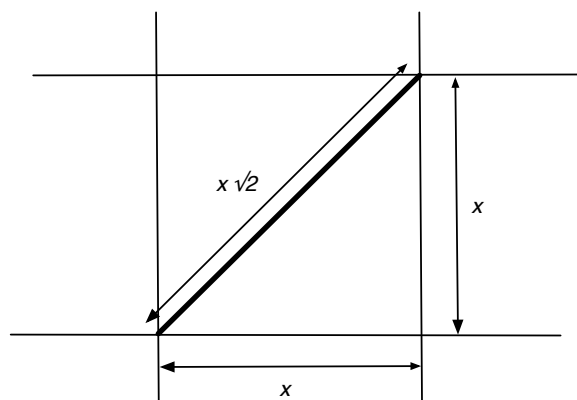


FIGURE 1 – Cas du déplacement sur un réseau carré : l'écart le plus important entre le trajet à vol d'oiseau et le trajet sur le réseau est obtenu pour une diagonale du réseau. Cet écart vaut  $2/\sqrt{2} = 1,414$ .

Christophe Terrier a étudié le problème de la mesure des longueurs des déplacements dans les enquêtes origine-destination dans une série d'article [Terrier 06, Terrier 09] qui sont à l'origine de ce travail. Le cœur de la question de [Terrier 09] est d'évaluer l'influence du maillage sur la mesure de flux entre zones, et des exercices de simulation ont permis de mettre en évidence deux phénomènes récurrents, quels que soient le type de maillage et la métrique utilisée (euclidienne ou Manhattan). Le premier est que, naturellement, le nombre de déplacements diminue lorsqu'on augmente la taille des mailles, puisque que de plus en plus de déplacements se trouvent être compris dans une seule maille, et donc ne sont plus captés dans la logique d'une enquête origine-destination (Figure 2). Le guide du CERTU préconise dans ce cas d'imputer les distances à  $\sqrt{\text{surface}_{\text{maille}}}/2$ .

Le deuxième est que, pour un grand nombre de voyages simulés, la somme des distances entre mailles est à peu près égale à la somme des distances entre points, avec des écarts relatifs compris entre 0,8 et 1,4, ce qui semble lié encore une fois à un effet d'orientation des mailles.

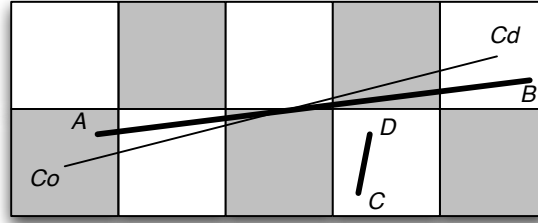


FIGURE 2 – Cas d’un maillage régulier : un déplacement entre les points  $A$  et  $B$  sera enregistré comme un déplacement entre la maille  $o$  origine de centroïde  $Co$  et la maille de destination  $d$  de centroïde  $Cd$ . La longueur du déplacement à vol d’oiseau sera  $d(Co, Cd)$  comme approximation de la vraie longueur à vol d’oiseau  $d(A, B)$ . Le déplacement entre  $C$  et  $D$  tombe à l’intérieur d’une maille.

Dans cette note, nous tentons d’aborder de manière formelle ces deux problèmes, en étudiant tout d’abord la probabilité qu’un déplacement soit interne à une maille, puis en étudiant la loi de probabilité de la distance entre deux points, selon la répartition spatiale du point de départ et du point d’arrivée. Dans la section 2, nous donnons une expression exacte de la probabilité qu’un déplacement soit interne à une maille carrée, en fonction de la taille de cette maille et de la longueur du déplacement. Dans la section 3, nous abordons le problème des maillages non réguliers en calculant la loi de la distance entre deux points tirés aléatoirement dans le voisinage de deux centres qui figurent les centres de deux mailles.

## 2 Cas d’un carroyage : quelques variations sur le problème de l’aiguille de Buffon

Avant d’entrer directement dans le vif du problème, nous faisons ici un rappel du problème de l’aiguille de Buffon [Buffon 33], qui nous servira dans la suite. En effet, comme relevé dans [Terrier 09], la détermination de la loi qui caractérise la probabilité qu’un déplacement soit inclus dans une maille est similaire au problème posé par Laplace en extension de l’expérience de l’aiguille de Buffon. Le lecteur intéressé par ces problèmes de probabilités géométriques pourra se référer au livre de Santaló [Santaló 04].

Rappelons que le problème dit de l’aiguille de Buffon est une expérience menée en 1733 par Georges-Louis Leclerc de Buffon et qui a permis une approximation du nombre  $\pi$ . Elle consiste en un lancer répété d’une aiguille de longueur  $l$  sur un parquet constitué de lattes parallèles de largeur  $a$ . Le quotient du nombre de fois où l’aiguille tombe à cheval sur deux lattes sur le nombre d’essais est une approximation de la probabilité que l’aiguille soit à cheval sur deux lattes lors d’un lancer et qui vaut exactement  $\frac{2l}{\pi a}$ .

La résolution initiale proposée par Buffon était purement géométrique : on remarque que les positions des aiguilles lors de  $n$  lancers forment un polygone à  $n$  côtés, et ce polygone s'approche d'un cercle lorsque  $n$  augmente. Laplace a généralisé ce résultat au cas d'un carrelage en utilisant le calcul intégral. Nous présentons brièvement ici la résolution du problème initial de Buffon, puis ses généralisations correspondant au problème posé par Jacques Choné dans [Terrier 09].

## 2.1 Cas d'un parquet (problème historique de Buffon)

On lance une aiguille de longueur  $l$  sur un parquet dont les lattes sont de largeur  $a$ . On suppose que  $a \geq l$ , c'est-à-dire que l'aiguille est plus petite que la largeur des lattes. Pour résoudre le problème, on note  $d$  la distance séparant le centre de l'aiguille à la rainure verticale la plus proche, et  $\theta$  l'angle compris entre  $-\pi/2$  et  $\pi/2$  qui mesure l'écart entre la direction de l'aiguille et une perpendiculaire aux lattes du parquet (Figure 3).

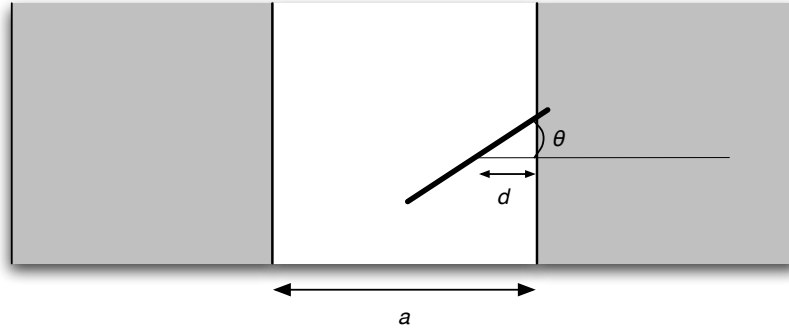


FIGURE 3 – Paramétrage du problème de l'aiguille de Buffon

Avec ce paramétrage, le lancer d'une aiguille est entièrement défini par le couple  $(\theta, d)$ , et l'aiguille est à cheval sur deux lattes si et seulement si

$$\frac{l \cos \theta}{2} \geq d$$

La probabilité recherchée est l'aire de la zone de chevauchement (zone favorable) divisée par l'aire de la zone totale, soit

$$\begin{aligned} \frac{\int_{-\pi/2}^{\pi/2} \int_0^{l \cos \theta / 2} d\theta dd}{\int_{-\pi/2}^{\pi/2} \int_0^{a/2} d\theta dd} &= \frac{\int_{-\pi/2}^{\pi/2} (l \cos \theta / 2) d\theta}{\pi a / 2} \\ &= \frac{l(\sin(\pi/2) - \sin(-\pi/2))/2}{\pi a / 2} \\ &= \frac{2l}{\pi a} \end{aligned}$$

qui est bien le résultat attendu. Le cas où  $a < l$  est plus délicat, et n'est pas pertinent dans le cadre de cet article. Nous passons maintenant au cas du problème de Loïc Terrier (95), qui reprend le problème de Laplace.

## 2.2 Cas d'un sol carrelé

Le problème 95 de Loïc Terrier cité dans [Terrier 09] est le suivant :

**Enoncé 1.** *On lance un spaghetti de longueur  $d$  sur un sol carrelé, les carreaux étant des carrés de côté unité. Quelle est la probabilité que le spaghetti soit à l'intérieur d'un des carreaux ?*

Le carrelage étant de côté 1, il est naturel d'étudier successivement les cas  $d \leq 1$  (aiguille plus petite que le côté d'un carreau, donc pouvant prendre n'importe quelle position à l'intérieur d'un carreau),  $1 \leq d \leq \sqrt{2}$  (aiguille plus petite que la diagonale de longueur  $\sqrt{2}$  du carreau) et  $d \geq \sqrt{2}$  (aiguille trop longue pour tenir à l'intérieur d'un carreau).

Nous utilisons un paramétrage différent de celui du problème de l'aiguille de Buffon. Nous repérons la position de l'aiguille par  $(x, y, \theta)$  avec  $x$  et  $y$  qui varient entre 0 et 1, et  $\theta \in [0, \pi/2]$  (Figure 4).

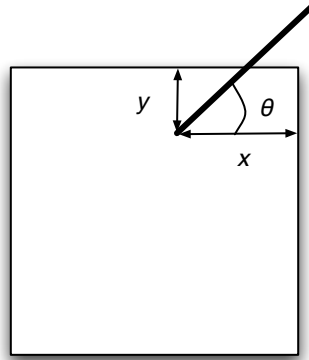


FIGURE 4 – Paramétrage du problème de Loïc Terrier

### 2.2.1 $d \leq 1$

Notons  $p$  la probabilité recherchée. Les cas défavorables sont ceux où l'aiguille touche une joint vertical, touche une joint horizontal, ou encore deux joints simultanément (Figure 5).

La probabilité  $p$  s'écrit en notant  $H$  l'événement «l'aiguille touche un joint horizontal» et  $V$  l'événement «l'aiguille touche un joint vertical» :

$$p = 1 - P(H \cap V)$$

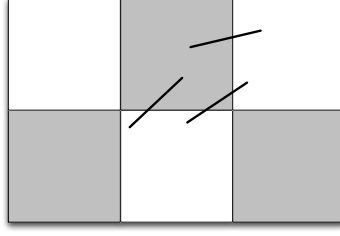


FIGURE 5 – Trois aiguilles de longueur  $d \leq 1$  représentant les trois cas défavorables : intersection d'un joint horizontal, intersection d'un joint vertical, intersection de deux joints

$$\begin{aligned}
 &= 1 - P(H) - P(V) + P(H \cup V) \\
 &= 1 - 2P(H) + P(H \cup V) \\
 &= 1 - 2p_1 + p_2
 \end{aligned}$$

Nous utilisons le résultat obtenu pour le problème de l'aiguille de Buffon puisque ce problème correspond à l'événement  $H$  et donc  $p_1 = 2d/\pi$ .

Le calcul de  $p_2$  est un tout petit plus compliqué. L'aiguille touche deux lames si on a à la fois  $x \leq d \cos \theta$  et  $y \leq d \sin \theta$ , avec  $\theta \in [0, \pi/2]$ , d'où :

$$\begin{aligned}
 p_2 &= \frac{\int_0^{\pi/2} \int_0^{x=d \cos \theta} \int_0^{y=d \sin \theta} dx dy d\theta}{\int_0^{\pi/2} \int_0^1 \int_0^1 dx dy d\theta} \\
 &= \frac{\int_0^{\pi/2} d^2 \cos \theta \sin \theta d\theta}{\pi/2} \\
 &= \frac{d^2}{\pi}
 \end{aligned}$$

Ainsi, on obtient que dans le cas où  $d \leq 1$ , la probabilité que l'aiguille (le spaghetti) soit à l'intérieur d'un des carreaux est

$$p = 1 - \frac{4d - d^2}{\pi}$$

.

### 2.2.2 $d \geq \sqrt{2}$

Comme l'aiguille est trop longue pour tenir à l'intérieur d'un carreau, la probabilité recherchée est nulle.

### 2.2.3 $1 \leq d \leq \sqrt{2}$

Ce cas est plus compliqué à résoudre. La probabilité recherchée s'écrit toujours  $p = 1 - 2p_1 + p_2$  mais  $p_1$  n'est pas simplement la probabilité du cas de l'aiguille de Buffon. Cette probabilité



correspond aux cas où  $y \leq d \sin \theta$  et  $y \leq 1$ , avec  $\theta \in [0, \pi/2]$ . Nous introduisons  $\alpha$  l'angle vérifiant  $d \sin \alpha = 1$ , soit  $\alpha = \arcsin(1/d)$ . Nous avons alors

$$\begin{aligned} p_1 &= \frac{\int_0^\alpha d \sin \theta d\theta + \int_\alpha^{\pi/2} d\theta}{\pi/2} \\ &= \frac{d - \sqrt{d^2 - 1} + \pi/2 - \arcsin(1/d)}{\pi/2} \end{aligned}$$

Le calcul de  $p_2$  nécessite un bon découpage de l'intégrale. Introduisons  $\beta = \arccos(1/d)$ . L'idée pour calculer  $p_2$  est de remarquer que

$$0 \leq \beta \leq \frac{\pi}{4} \leq \alpha \leq \frac{\pi}{2}$$

et ainsi on obtient que

$$\begin{aligned} p_2 &= \frac{\int_0^\beta d \sin \theta d\theta + \int_\beta^{\pi/4} d^2 \sin \theta \cos \theta d\theta + \int_{\pi/4}^\alpha d^2 \cos \theta \sin \theta d\theta + \int_\alpha^{\pi/2} d \cos \theta d\theta}{\pi/2} \\ &= \frac{d(1 - 1/d) + d^2/2(1/2 - (1 - 1/d^2)) + d^2/2(1/d^2 - 1/2)) + d(1 - 1/d)}{\pi/2} \\ &= \frac{4d - 2 - d^2}{\pi} \end{aligned}$$

La probabilité recherchée  $p$  vaut donc

$$p = 1 - \frac{d^2 - 4\sqrt{d^2 - 1} - 4 \arcsin(1/d) + 2\pi + 2}{\pi}$$

La probabilité limite pour  $d = 1$  vaut  $p = 1 - \frac{3}{\pi}$ .

Dans le cas d'un maillage régulier carré de pas unité, la probabilité qu'un trajet de longueur  $d$  soit interne à un carreau est une fonction décroissante en  $d$  (Figure 6), que nous avons déterminée complètement :

$$\begin{aligned} p &= 1 - \frac{4d - d^2}{\pi} \text{ pour } d \leq 1 \\ &= 1 - \frac{d^2 - 4\sqrt{d^2 - 1} - 4 \arcsin(1/d) + 2\pi + 2}{\pi} \text{ pour } 1 \leq d \leq \sqrt{2} \\ &= 0 \text{ pour } d \geq \sqrt{2} \end{aligned}$$

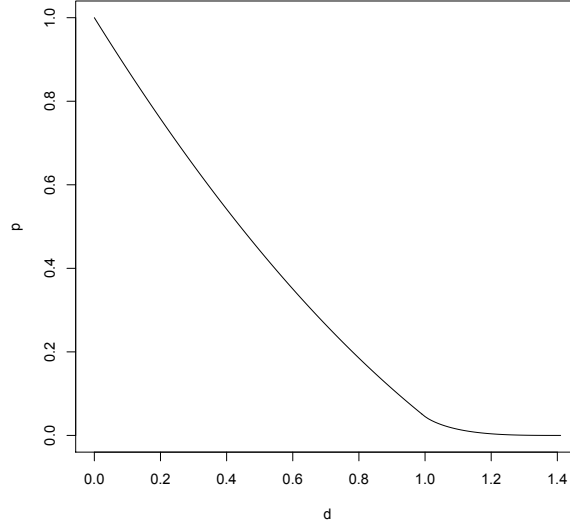


FIGURE 6 – Probabilité qu’un trajet de longueur  $d$  soit compris entièrement dans un carré de côté unité en fonction de  $d$

## 2.3 Cas d’une grande aiguille sur un sol carrelé

Nous revenons au sujet des longueurs des trajets calculés à partir des enquêtes origine-destination en prenant le cas d’une grande aiguille, figurant le trajet, sur un carroyage décrivant les zones d’enquêtes tel que celui de l’EGT. Pour simplifier les calculs, nous revenons au cas simple du parquet, mais les résultats s’étendent de la même façon que plus haut au cas d’un dallage à partir des probabilités que nous avons calculées.

On appelle  $X$  la variable aléatoire qui décrit le nombre de lames intersectées par l’aiguille lors d’un lancer. Nous nous intéressons à l’espérance mathématique de  $X$  notée  $E(X)$ . Cette espérance est une fonction de la longueur de l’aiguille  $d$ , et nous notons cette fonction  $f(d) = E(X)$ .

### 2.3.1 Cas où $d \leq a$

La variable aléatoire est une variable de Bernoulli de paramètre  $p = \frac{2d}{\pi a}$  donc

$$E(X) = \frac{2d}{\pi a} = f(d)$$

### 2.3.2 Cas où $d \geq a$

L'espérance étant une fonction linéaire,  $f(d)$  est clairement linéaire en  $d$  et on écrit  $d = na + d_0$  avec  $d_0 \leq a$  et  $n \in \mathbb{N}$ , et on développe ainsi :

$$f(d) = f(na + d_0) = nf(a) + f(d_0) = n\frac{2a}{\pi a} + \frac{2d_0}{\pi a} = \frac{2d}{\pi a}$$

On obtient donc que, quelque soit la longueur de l'aiguille, on a en moyenne  $\frac{2d}{\pi a}$  lames touchées. Ce résultat est valable dans le cas où l'aiguille est un segment de droite, mais également dans le cas où elle est une ligne polygonale.

## 3 Cas d'un maillage quelconque

Nous abordons maintenant le deuxième problème, qui est de déterminer la lois des distances entre un deux trajets, en fonction de la distance entre la maille origine et la maille destination.

À partir d'un maillage quelconque, nous supposons qu'un trajet a été effectué entre deux points  $A$  et  $B$ , où  $A$  appartient à la zone de départ  $o$  de centroïde  $C_o = A_0$  et  $B$  à la zone d'arrivée de centroïde  $C_d = B_0$ . La vraie distance du parcours est  $d(A, B)$ , et la distance imputée à vol d'oiseau  $d(C_o, C_d) = d(A_0, B_0)$ .

### 3.1 Expression générale

En plaçant un repère orthonormé dont l'origine est en  $A_0$  et un des axes de même direction que  $\overrightarrow{A_0 B_0}$  on a dans ce repère  $A_0 = (0, 0)$ ,  $B_0 = (d_0, 0)$ ,  $A = (x_A, y_A)$  et  $B = (x_B, y_B)$ . On peut supposer que  $x_A, x_B, y_A$  et  $y_B$  sont des réalisations de variables aléatoire indépendantes  $X_A, X_B, Y_A$  et  $Y_B$ .

Avec les notation précédentes, on peut chercher à déterminer directement la loi de  $Z = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$  qui est la distance (aléatoire) entre  $A$  et  $B$ . Comme  $X_A, X_B, Y_A$  et  $Y_B$  sont supposées indépendantes de densités  $f_A(), f_B(), g_A()$  et  $g_B()$ , la densité de  $X = X_A - X_B$  est  $f(x) = \int f_A(v)f_B(v-x)dv$  et la densité de  $Y = Y_A - Y_B$  est  $g(y) = \int g_A(v)g_B(v-y)dv$ .

On en déduit la densité de  $Z$  qui est norme euclidienne du couple  $(X, Y)$  en calculant pour un fonction  $h$  :

$$E(h(Z)) = \int \int h(\sqrt{x^2 + y^2})f(x)g(y)dx dy$$

qui se calcule en effectuant un changement de variables en polaires :

$$\begin{aligned} E(h(Z)) &= \int \int h(r)r f(r \cos \theta)g(r \sin \theta)dr d\theta \\ &= \int_0^\infty h(r) \int_0^{2\pi} r f(r \cos \theta)g(r \sin \theta)d\theta dr \end{aligned}$$

ce qui donne immédiatement que la densité de  $Z$  est :

$$h(r) = r \int_0^{2\pi} f(r \cos \theta) g(r \sin \theta) d\theta$$

soit en réintroduisant les densités de  $X_A$ ,  $X_B$ ,  $Y_A$  et  $Y_B$  :

$$h(r) = r \int_0^{2\pi} \left( \int f_A(v) f_B(v - r \cos \theta) dv \int g_A(v) g_B(v - r \sin \theta) dv \right) d\theta$$

Cette densité est calculable numériquement quelque soient les densités des variables décrivant la position de  $A$  et de  $B$ .

### 3.2 Illustrations

Nous présentons ici quelques applications numériques pour illustrer le calcul précédent. Nous considérons un trajet entre deux mailles dont les centroïdes sont distants de 60km :  $d^2(A_0, B_0) = d_0 = 60$ . Nous supposons dans un premier temps que  $A_0$  et  $B_0$  sont des chefs-lieux, et donc que  $X_A$  et  $Y_A$  sont des lois normales indépendantes de même variance, de même que  $X_B$  et  $Y_B$ .

La figure 7 présente la loi des distances pour les distributions normales centrées suivantes :

$$\begin{aligned} X_A &\sim N(0, \sigma_A^2) \\ Y_A &\sim N(0, \sigma_A^2) \\ X_B &\sim N(60, \sigma_B^2) \\ Y_B &\sim N(0, \sigma_B^2) \end{aligned}$$

avec  $\sigma_A^2 = \sigma_B^2 = 9$ , puis  $\sigma_A^2 = 1$  et  $\sigma_B^2 = 9$ .

Le cas où les distributions ont même variance donne une distribution pour  $h$  très proche de celle d'une loi normale  $N(60, 18)$ , ce qui était attendu. En effet, la distance :

$$d(A, B) = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$$

a pour terme principal  $|X_A - X_B|$ .

## 4 Cas d'un maillage carré

Nous considérons maintenant un maillage carré avec deux mailles sur une même ligne horizontale séparées d'une distance centre à centre de 60km, et deux mailles séparées par une même distance en diagonale (Figure 8). Ces deux cas ne peuvent appartenir naturellement au même maillage, mais nous les rapprochons pour mettre en évidence l'effet de l'orientation des mailles.

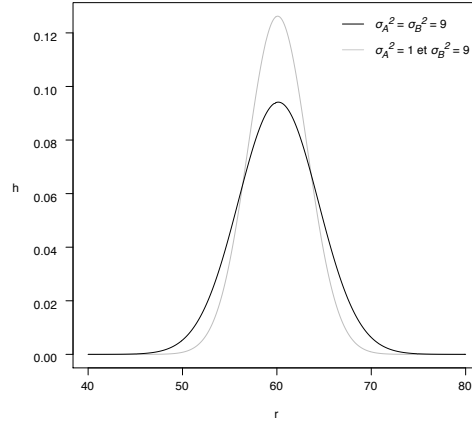


FIGURE 7 – Distribution des distances pour un trajet dont l’origine et la destination sont réparties de manière normale autour de deux points distants de 60km

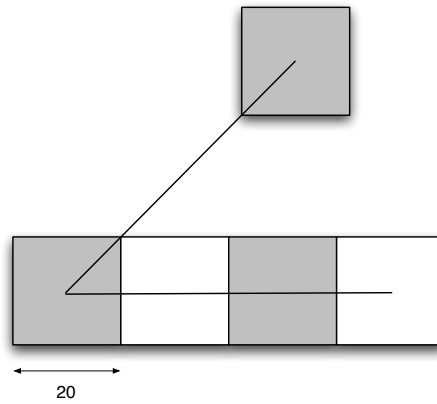


FIGURE 8 – Deux mailles séparées de 60km à l’horizontale, et en diagonale

Les lois des positions des origines et des destinations sont alors les suivantes :

$$X_A \sim U[-10, 10]$$

$$Y_A \sim U[-10, 10]$$

$$X_B \sim U[50, 70]$$

$$Y_B \sim U[-10, 10]$$

pour les mailles alignées, et

$$X_A \sim U[-10, 10]$$

$$\begin{aligned}
Y_A &\sim U[-10, 10] \\
X_B &\sim U[\sqrt{60}/2 - 10, \sqrt{60}/2 + 10] \\
Y_B &\sim U[\sqrt{60}/2 - 10, \sqrt{60}/2 + 10]
\end{aligned}$$

pour les mailles en diagonale.

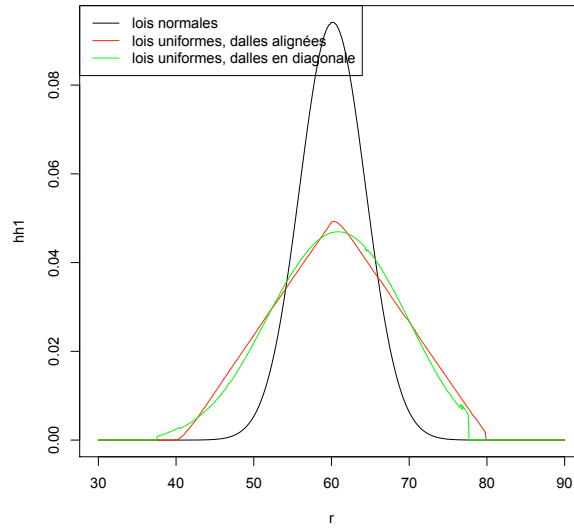


FIGURE 9 – Lois des distances pour des trajets dont les origines sont réparties uniformément dans des mailles distantes de 60km à l’horizontale (rouge), et en diagonale (vert)

Les approximations numériques perturbent un peu le tracé des courbes, mais on constate (Figure 9) l’effet de l’orientation des mailles sur les lois des distances.

Ici la distance  $d(A, B)$  a encore pour terme principal  $|X_A - X_B|$  dont un raisonnement géométrique montre qu’il suit une loi triangulaire sur  $[40, 80]$  de densité :

$$\begin{aligned}
f(r) &= 0 && \text{si } r \leq 40 \\
&= \frac{r-40}{400} && \text{si } 40 \leq r \leq 60 \\
&= \frac{80-r}{400} && \text{si } 60 \leq r \leq 80 \\
&= 0 && \text{si } r \geq 80
\end{aligned}$$

lorsque les carrés sont alignés.

## 4.1 Cas d'un maillage carré, avec agglomération autour du lieu de départ

Enfin, nous donnons comme illustration de notre loi le cas où, toujours sur un maillage carré, les origines sont concentrées autour du centre de la maille. On utilise alors une loi normale (qui devrait être tronquée, en toute rigueur) pour l'origine, et une loi uniforme pour la destination :

$$X_A \sim U[-10, 10]$$

$$Y_A \sim U[-10, 10]$$

$$X_B \sim N(60, 9)$$

$$Y_B \sim N(0, 9)$$

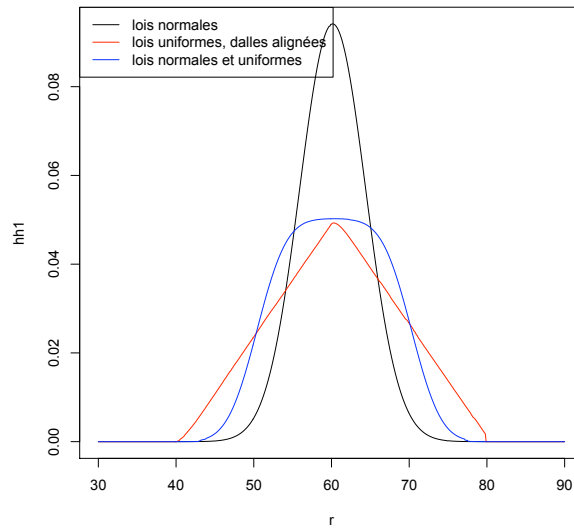


FIGURE 10 – Lois des distances pour des trajets dont les origines sont réparties de manière normale (noir), uniformément dans des mailles distantes de 60km à l'horizontale (rouge), et de manière normale à l'origine et uniforme à la destination (bleu)

La densité obtenue est présentée Figure 10. Nous observons un très fort étalement du mode de la distribution, correspondant au fait que le point destination est réparti uniformément dans la maille de côté 20.

## Références

- [Buffon 33] Buffon, G.-L. (1733). *Mémoire sur le jeu du Franc Carreau*, présenté à l'Académie des Sciences.
- [CERTU] CERTU. *Calcul a posteriori des distances dans les enquêtes ménages déplacements*, note méthodologique du CERTU.
- [Santaló 04] Santaló, L. (2004). *Integral Geometry and Geometric Probability*. Cambridge University press.
- [Terrier 06] Terrier, C. (2006). Flux et afflux de touristes : les instruments de mesure, la géomatématique des flux, *FLUX, Cahiers scientifique internationaux Réseaux et Territoires*, **65**.
- [Terrier 09] Terrier, C. (2009). Flux et afflux des voyageurs : géomatématique des flux et procédés d'enquête, *Journées de méthodologie statistique de l'INSEE 2009*.